

University of Groningen

Limit theory for the sample autocorrelations and extremes of a GARCH (1,1) process

Mikosch, T; Starica, C

Published in:
Annals of statistics

DOI:
[10.1214/aos/1015957401](https://doi.org/10.1214/aos/1015957401)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2000

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Mikosch, T., & Starica, C. (2000). Limit theory for the sample autocorrelations and extremes of a GARCH (1,1) process. *Annals of statistics*, 28(5), 1427-1451. <https://doi.org/10.1214/aos/1015957401>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

LIMIT THEORY FOR THE SAMPLE AUTOCORRELATIONS AND EXTREMES OF A GARCH (1, 1) PROCESS¹

BY THOMAS MIKOSCH AND CĂTĂLIN STĂRICĂ

University of Groningen and Chalmers University of Technology

The asymptotic theory for the sample autocorrelations and extremes of a GARCH(1, 1) process is provided. Special attention is given to the case when the sum of the ARCH and GARCH parameters is close to 1, that is, when one is close to an infinite variance marginal distribution. This situation has been observed for various financial log-return series and led to the introduction of the IGARCH model. In such a situation, the sample autocorrelations are unreliable estimators of their deterministic counterparts for the time series and its absolute values, and the sample autocorrelations of the squared time series have nondegenerate limit distributions. We discuss the consequences for a foreign exchange rate series.

1. Introduction. Log-returns $X_t = \ln P_t - \ln P_{t-1}$ of foreign exchange rates, stock indices and share prices P_t , $t = 1, 2, \dots$, typically share the following features:

1. The frequency of large and small values (relative to the range of the data) is rather high, suggesting that *the data do not come from a normal, but from a heavy-tailed distribution*.
2. Exceedances of high thresholds occur in clusters, which indicates that *there is dependence in the tails*.
3. Sample autocorrelations of the data are tiny whereas the sample autocorrelations of the absolute and squared values are significantly different from zero even for large lags. This behavior suggests that *there is some kind of long-range dependence in the data*.

Various models have been proposed in order to describe these features. Among them, models of the type

$$X_t = \sigma_t Z_t, \quad t \in \mathbb{Z},$$

have become particularly popular. Here (Z_t) is a sequence of iid symmetric random variables with $EZ_1^2 = 1$. One often assumes the Z_t 's to be standard normal. Moreover, the sequence (σ_t) consists of nonnegative random variables such that Z_t and σ_t are independent for every fixed t . We frequently refer to σ_t as the *stochastic volatility* of X_t . Models of this type include the ARCH and GARCH family; see for example [16] for their definitions and properties.

Received December 1998; revised April 2000.

¹Supported in part by Chalmers University and the University of Groningen.

AMS 1991 *subject classifications*. Primary 62P20; secondary 90A20, 60G55, 60J10, 62F10, 62F12.

Key words and phrases. GARCH, sample autocorrelations, stochastic recurrence equation, Pareto tail, extremes, extremal index, point processes, foreign exchange rates.

We often write σ for a generic random variable with the distribution of σ_1 , X for a generic random variable with the distribution of X_1 , etc.

We restrict ourselves to one particular model which has very often been used in applications: the GARCH(1, 1) process. It is defined by specifying σ_t as follows:

$$\sigma_t^2 = \alpha_0 + \beta_1 \sigma_{t-1}^2 + \alpha_1 X_{t-1}^2 = \alpha_0 + \sigma_{t-1}^2 (\beta_1 + \alpha_1 Z_{t-1}^2), \quad t \in \mathbb{Z}.$$

The parameters α_0 , α_1 and β_1 are nonnegative.

The stationary GARCH(1, 1) process is believed to capture, despite its simplicity, various of the empirically observed properties of log-returns. (Stationarity is always understood as strict stationarity.) For example, the stationary GARCH(1, 1) processes can exhibit heavy-tailed marginal distributions of power law type and hence they could be appropriate tools to model the heavier-than-normal tails of the financial data. This follows from a classical result by Kesten [29]; see Theorem 2.1 below. Although this result does not seem to be well known in the econometrics literature, the fact that certain power moments of X need not exist has been known for a long time; see, for example, [37]. The question about the extent to which the tails of the estimated GARCH(1, 1) model do describe the tails of the empirical distributions was addressed in [44]. It is shown there that, when using normal innovations, the tails of the fitted GARCH(1, 1) models seem to be much thinner than the tails apparent in the data. Hence, even though the GARCH(1, 1) processes could display heavy tails, when estimated on the data they *do not* produce tails that match the empirical ones. The relationship between the tail index of a GARCH(1, 1) process, its coefficients and the distribution of the innovations is made clear in Section 2.2.

The tail behavior of the fitted GARCH(1, 1) processes is also important from another perspective. The empirical fact that the GARCH(1, 1) models fitted to log-return data often satisfy the condition $\alpha_1 + \beta_1 \approx 1$ implies that one often deals with a class of models with $E|X|^{2+\delta} = \infty$ for δ close to zero. For such models, the asymptotic behavior of various classical time series tools such as the sample autocorrelations and the periodogram are not always well understood and give rise to many theoretical questions. [The GARCH(1, 1) model with $\alpha_1 + \beta_1 = 1$ is called integrated GARCH(1, 1) or IGARCH(1, 1); see [17].]

Another empirical finding concerns the behavior of the sample autocorrelation function (sample ACF) of powers of absolute log-return data at large lags. It has been noticed that the mentioned sample autocorrelations decay to zero at a hyperbolic rate ("long-range dependence"). This seems to be in contradiction with the sample ACF behavior of GARCH(1, 1) model. The GARCH(1, 1) process has good mixing properties; it is strongly mixing with geometric rate, provided Z has a density and $E|Z|^\varepsilon < \infty$ for some $\varepsilon > 0$; see for example [12]. Hence the autocorrelations of the underlying process, its absolute values and squares, given these quantities are well defined and decrease to zero at an exponential rate.

However, as mentioned above, most often the fitted GARCH(1, 1) models for log-return data belong to the class of GARCH(1, 1) processes with very heavy tails, that is, models which do not have a finite fourth moment, although their second moment may still exist. Hence, autocovariances and autocorrelations are either not defined (for the squares, third powers, etc.), or when they exist (for the time series itself and its absolute values) the standard theory for the sample autocorrelation function, that is, Gaussian limit distributions and \sqrt{n} -rates of convergence, is not valid any more. We show that in these cases the sample autocorrelations have infinite variance distributional limits and the rates of convergence are extremely slow. As a result, the asymptotic confidence bands are much wider than in the classical asymptotically normal theory. The fact that the sample ACF of ARCH-type processes has wider confidence bands than for linear processes has already been mentioned in [4].

Under these circumstances one could hope that the confidence bands are perhaps wide enough to bound the apparently hyperbolically decaying sample autocorrelation function of the absolute values of log-returns. In other words, it is possible that the discrepancy we mentioned between the empirically observed hyperbolic decay rate in the sample autocorrelation function and the exponential decay of the autocorrelation function of the GARCH(1, 1) model could be explained through statistical uncertainty related to the estimation procedure and hence claimed to be insignificant. If this were true, then, up to statistical uncertainty, the GARCH(1, 1) model could be said to explain at least the sample autocorrelation function behavior. One of the conclusions of the paper is that, even when the mentioned larger-than-usual statistical uncertainty is accounted for, the GARCH(1, 1) *cannot* explain the effect of almost constancy of the sample ACF of the absolute values of log-returns. The latter phenomenon can be explained by the nonstationary of the data; see [35] for an extensive discussion.

As another desirable property that would recommend the GARCH(1, 1) model as a viable candidate that captures the already mentioned common features of the financial log-returns, exceedances of very low–high thresholds by the GARCH(1, 1) process tend to occur in clusters. Formally, this behavior can be described by the weak convergence of the point processes of exceedances, associated with the time series, to a compound Poisson process. The cluster sizes of this limiting process determine the *extremal index* $\theta \in (0, 1)$, $1/\theta$ being the expected size of the clusters. Section 4 is devoted to the extremal behavior of the GARCH(1, 1). A comparison of the estimated extremal indices of simulated GARCH(1, 1) and foreign exchange rate (FX) data is given in Section 6. This analysis reveals that the GARCH(1, 1) model fit to the log-returns does not, once again, properly describe the observed features of the data. The expected cluster sizes of the exceedances of high–low thresholds of the FX log-returns are smaller than the expected cluster sizes of simulated GARCH(1, 1) processes whose parameters were estimated from the FX observations. This means that there is less dependence in the tails of real-life data than in the GARCH(1, 1) model.

Our results serve in our view a double goal. On the one hand, they can be thought of as a tool for deciding to which extent the potentially useful features of the GARCH(1, 1) model (heavy tails, slowly decaying sample ACF in the case $\alpha_1 + \beta_1 \approx 1$, clustering of the extremes) do actually describe accurately the corresponding empirical behavior. In this sense, we conclude that, although displaying useful features, the GARCH(1, 1) model does not seem to accurately describe either the extremal behavior or the correlation structure captured by the sample ACF of the data set that we analyzed in detail.

On the other hand, we think of our findings as contributions to the growing number of results that emphasize the serious differences between the behavior of various statistical tools under light and heavy tails when dependency is present; see [41] for a recent survey paper. In this direction, we showed that the sample ACFs of GARCH(1, 1) models, their absolute values, squares, third powers, etc., fitted to real-life FX log-returns, are either poor estimators of the ACFs (slow convergence rates) or meaningless (nondegenerate limit distributions). Hence, in the case of the GARCH(1, 1) modelling, the sample ACF can be an extremely problematic statistical instrument that has to be used with caution when making statistical statements.

The paper is organized as follows. In Section 2 we consider some basic theoretical properties of the GARCH(1, 1) model. The weak convergence of the point processes associated with the sequences (X_t, σ_t) , $(|X_t|, \sigma_t)$ and (X_t^2, σ_t^2) is considered in Section 3. In Section 4 we use these results to study the extremal behavior of a GARCH(1, 1) process, including the calculation of its extremal index, the weak convergence of the point processes of exceedances and the weak limits of the distributions of the extremes. In Section 5 we study the asymptotic behavior of the sample autocovariances and autocorrelations of the σ_t 's and X_t 's, their squares and absolute values. Section 6 contains an empirical study of foreign exchange rates and simulated GARCH(1, 1). In particular, we check the appropriateness of the GARCH(1, 1) as a model for the observed data as regards their dependence structure described by the autocorrelation and autocovariance functions, tails and extremal behavior.

We conclude this section by noting that, for the sake of conciseness and due to the strong connection between the present work and Davis and Mikosch [11], we use notation and results from [11] without explicitly stating them. For easy reference an extended, self-contained version of this paper is available via the Internet; see [34].

2. Basic properties of GARCH(1, 1). In what follows, we collect some facts about the probabilistic properties of the GARCH(1, 1). First we notice that the GARCH(1, 1) can be considered in the much wider context of stochastic recurrence (or difference) equations of type

$$(2.1) \quad \mathbf{X}_t = \mathbf{A}_t \mathbf{X}_{t-1} + \mathbf{B}_t, \quad t \in \mathbb{Z},$$

where $((\mathbf{A}_t, \mathbf{B}_t))$ is an iid sequence, for every t the vector \mathbf{X}_{t-1} is independent of $(\mathbf{A}_t, \mathbf{B}_t)$, the \mathbf{A}_t 's are iid random $d \times d$ matrices and the \mathbf{B}_t 's are iid d -dimensional random vectors.

Indeed, write

$$(2.2) \quad \mathbf{X}_t = \begin{pmatrix} X_t^2 \\ \sigma_t^2 \end{pmatrix}, \quad \mathbf{A}_t = \begin{pmatrix} \alpha_1 Z_t^2 & \beta_1 Z_t^2 \\ \alpha_1 & \beta_1 \end{pmatrix}, \quad \mathbf{B}_t = \begin{pmatrix} \alpha_0 Z_t^2 \\ \alpha_0 \end{pmatrix}.$$

Then \mathbf{X}_t satisfies equation (2.1) with $d = 2$. Also observe that σ_t^2 satisfies the recurrence equation

$$(2.3) \quad \sigma_t^2 = \alpha_0 + \sigma_{t-1}^2(\alpha_1 Z_{t-1}^2 + \beta_1), \quad t \in \mathbb{Z},$$

which is of the same type as (2.1) for $d = 1$, with $\mathbf{X}_t = \sigma_t^2$, $\mathbf{A}_t = \alpha_1 Z_{t-1}^2 + \beta_1$ and $\mathbf{B}_t = \alpha_0$.

Equations of type (2.1) have been extensively studied; see [1, 5, 6, 7, 19, 20, 13, 29, 45] and the references therein.

2.1. Existence of a stationary solution. The first question regards the existence of a stationary non-anticipative solution to equations (2.1) and (2.3). Applying the results of [5] (the latter paper contains the most complete results; for related work see [7, 45] in the case $d = 1$ and [19] for $d \geq 1$) it follows that (2.1) has a stationary solution if $E \ln^+ \|\mathbf{A}\| < \infty$, $E \ln^+ |\mathbf{B}| < \infty$ and if the top Lyapunov exponent $\tilde{\gamma}$ defined as $\tilde{\gamma} = \inf \{n^{-1} E \ln \|\mathbf{A}_1 \cdots \mathbf{A}_n\|, n \in \mathbb{N}\}$, is negative. Here $|\cdot|$ is any norm in \mathbb{R}^n , and $\|\mathbf{A}\| = \sup_{|\mathbf{x}|=1} |\mathbf{A}\mathbf{x}|$ is the corresponding operator norm. Moreover, these conditions are close to necessity; see [5]. In [6] they also studied the stationarity of the squared stochastic volatility process σ_t^2 for a general GARCH (p, q) process; the case $p = q = 1$ was treated in [37].

It is in general difficult to calculate the top Lyapunov exponent $\tilde{\gamma}$. However, in the particular case (2.2) calculation yields

$$(2.4) \quad \mathbf{A}_n \cdots \mathbf{A}_1 = \mathbf{A}_n \prod_{t=1}^{n-1} (\alpha_1 Z_t^2 + \beta_1),$$

and so $\tilde{\gamma} = E \ln(\alpha_1 Z^2 + \beta_1)$, provided $E |\ln |Z|| < \infty$. Alternatively, one can use the one-dimensional equation (2.3) and conclude from the conditions and literature above that

$$(2.5) \quad \alpha_0 > 0 \quad \text{and} \quad E \ln(\alpha_1 Z^2 + \beta_1) < 0$$

are necessary and sufficient for stationarity of (X_t^2, σ_t^2) ; see [1] for details. Moreover, $\beta_1 < 1$ is necessary; see [6]. Notice that stationarity of σ_t^2 implies stationarity of the sequence $(X_t^2, \sigma_t^2) = \sigma_t^2 (Z_t^2, 1)$, $t \in \mathbb{Z}$. By construction of the sequence (X_t) , stationarity of the sequence (X_t, σ_t) follows. *In what follows, we assume that condition (2.5) is satisfied.* Then a stationary version of (X_t, σ_t) exists.

2.2. The tails of X_t and σ_t . In this section we consider the tail behavior of the vectors (X_t, σ_t) which are the basic building blocks for the point processes of Section 3. It can be described by multivariate regular variation; see [39, 40] for properties and applications of this notion. A d -dimensional vector \mathbf{Y} is

said to be *regularly varying with index* $\kappa \geq 0$ and *spectral measure* P_Θ on the Borel σ -field of the unit sphere \mathbb{S}^{d-1} of \mathbb{R}^d if

$$(2.6) \quad \frac{P(|\mathbf{Y}| > xt, \mathbf{Y}/|\mathbf{Y}| \in \cdot)}{P(|\mathbf{Y}| > t)} \xrightarrow{v} x^{-\kappa} P_\Theta(\cdot).$$

Here \rightarrow^v denotes vague convergence (see [28] for its definition) and P_Θ is the distribution of a certain random vector Θ with values in \mathbb{S}^{d-1} .

The following result is a consequence of the renewal theory for products of random matrices due to [29]. By virtue of the special structure of the GARCH(1, 1) process one can, however, reduce the problem to one-dimensional renewal theory in which case one can apply the elegant results of [20]. For convenience, we write

$$A_t = \alpha_1 Z_{t-1}^2 + \beta_1 \quad t \in \mathbb{Z}.$$

THEOREM 2.1. *Assume the law of $\ln A$ is nonarithmetic, $E \ln A < 0$, $P(A > 1) > 0$ and there exists $h_0 \leq \infty$ such that $EA^h < \infty$ for all $h < h_0$ and $EA^{h_0} = \infty$. Then the following statements hold:*

(a) *The equation*

$$(2.7) \quad EA^{\kappa/2} = 1$$

has a unique positive solution.

(b) *Assume $\alpha_0 > 0$ and κ satisfies (2.7). Then there exists a stationary solution (σ_t^2) to (2.3). For independent A, σ^2 with $A \stackrel{d}{=} A_1$ and $\sigma^2 \stackrel{d}{=} \sigma_1^2$, there exists a positive constant $c_0 = E[(\alpha_0 + A\sigma^2)^{\kappa/2} - (A\sigma^2)^{\kappa/2}]/[(\kappa/2)EA^{\kappa/2} \ln A]$ such that*

$$(2.8) \quad P(\sigma > x) \sim c_0 x^{-\kappa}$$

and

$$P(|X| > x) \sim E|Z|^\kappa P(\sigma > x) \quad \text{as } x \rightarrow \infty.$$

Moreover, the vector (X, σ) is jointly regularly varying with index κ and spectral measure on \mathbb{S}^1 given by

$$P(\Theta \in \cdot) = \frac{E|(Z, 1)|^\kappa I_{\{(Z, 1)/|(Z, 1)| \in \cdot\}}}{E|(Z, 1)|^\kappa}.$$

REMARK 2.2. If a unique $\kappa > 0$ with (2.7) exists, $-\infty \leq E \ln A < 0$ and $P(A > 1) > 0$ hold necessarily; see [46]. Since $\beta_1 < 1$ is necessary for stationarity, $\alpha_1 = 0$ is not a possible parameter choice when (2.7) holds.

If the conditions of Theorem 2.1 hold and $\alpha_1 + \beta_1 = 1$ then (2.7) has the unique solution $\kappa = 2$. This implies that $P(|X| > x) \sim cx^{-2}$ for some $c > 0$ and, in turn, that $EX^2 = \infty$. GARCH(1, 1) models fitted to log-returns frequently have parameters α_1 and β_1 such that $\alpha_1 + \beta_1$ is close to 1. This indicates that one deals with time series models with extremely heavy tails.

PROOF. The function EA^h is continuous and convex in h . Since $E \ln A < 0$ it assumes values smaller than 1 in some neighborhood of the origin. Moreover, for sufficiently large h , since $P(A > 1) > 0$ and $EA^{h_0} = \infty$, $EA^h \geq 1$.

Therefore $EA^{\kappa/2} = 1$ has a unique positive solution. Since condition (2.5) holds, a stationary solution (σ_t^2) exists. The assumptions ensure that $EA^{\kappa/2+\varepsilon} < \infty$ for small $\varepsilon > 0$ and so all conditions of Theorem 4.1 in [20] are satisfied. The latter result gives the first relation in (2.8). The second one is a consequence of a result by Breiman [8]: assume ξ and η are independent nonnegative random variables such that $P(\xi > x) = L(x)x^{-\kappa}$ for some slowly varying function L and $E\eta^{\kappa+\varepsilon} < \infty$ for some $\varepsilon > 0$. Then

$$(2.9) \quad P(\eta\xi > x) \sim E\eta^\kappa P(\xi > x), \quad x \rightarrow \infty.$$

Another application of Breiman's result yields for any Borel set $B \subset \mathbb{S}^1$ that

$$\begin{aligned} P(|(X, \sigma)| > xt, (X, \sigma)/|(X, \sigma)| \in B) &= P(\sigma|(Z, 1)|I_{\{(Z, 1)/|(Z, 1)| \in B\}} > xt) \\ &\sim E|(Z, 1)|^\kappa I_{\{(Z, 1)/|(Z, 1)| \in B\}} x^{-\kappa} P(\sigma > t), \\ P(|(X, \sigma)| > t) &= P(\sigma|(Z, 1)| > t) \sim E|(Z, 1)|^\kappa P(\sigma > t). \end{aligned}$$

This concludes the proof. \square

The idea of the proof of Theorem 2.1 can be used to derive the joint tail behavior of the following lagged vectors, for $h \geq 0$,

$$\begin{aligned} \mathbf{Y}_h &= (X_0, \sigma_0, \dots, X_h, \sigma_h), \\ (2.10) \quad \mathbf{Y}_h^{(i)} &= (|X_0|^i, \sigma_0^i, \dots, |X_h|^i, \sigma_h^i), \quad i = 1, 2, \\ \mathbf{Z}_h^{(2)} &= \left((Z_0^2, 1), A_1(Z_1^2, 1), \prod_{i=1}^2 A_i(Z_2^2, 1), \dots, \prod_{i=1}^h A_i(Z_h^2, 1) \right). \end{aligned}$$

This particular form of vectors is used in Section 5 to construct the sample ACF of the time series (X_t) , its absolute values and squares.

THEOREM 2.3. *Let $h \geq 0$ and assume that the conditions of the Theorem 2.1(b) are satisfied. Let $|\cdot|$ denote the max-norm.*

(a) $\mathbf{Y}_h^{(2)}$ is regularly varying with index $\kappa/2$ and spectral measure

$$P_\Theta(\cdot) = \frac{E|\mathbf{Z}_h^{(2)}|^{\kappa/2} I_{\{\mathbf{Z}_h^{(2)}/|\mathbf{Z}_h^{(2)}| \in \cdot\}}}{E|\mathbf{Z}_h^{(2)}|^{\kappa/2}},$$

where $\Theta = (\theta_{0, X^2}, \theta_{0, \sigma^2}, \dots, \theta_{h, X^2}, \theta_{h, \sigma^2})$.

(b) $\mathbf{Y}_h^{(1)}$ is regularly varying with index κ and spectral measure $P_{\Theta^{1/2}}$, where $\Theta^{1/2}$ is obtained from Θ by componentwise taking square roots.

(c) \mathbf{Y}_h is regularly varying with index κ and spectral measure given by the distribution of the vector

$$(2.11) \quad (r_0 \theta_{0, X^2}^{1,2}, \theta_{0, \sigma^2}^{1/2}, \dots, r_h \theta_{h, X^2}^{1/2}, \theta_{h, \sigma^2}^{1/2}),$$

where (r_t) is a sequence of iid Bernoulli random variables such that $P(r = \pm 1) = 0.5$, independent of Θ .

PROOF. We start with

$$\begin{aligned}\mathbf{Y}_h^{(2)} &= \left(\sigma_0^2(Z_0^2, 1), \sigma_1^2(Z_1^2, 1), \dots, \sigma_h^2(Z_h^2, 1) \right) \\ &= \left(\sigma_0^2(Z_0^2, 1), (\alpha_0 + \sigma_0^2 A_1)(Z_1^2, 1), \dots, (\alpha_0 + \sigma_{h-1}^2 A_h)(Z_h^2, 1) \right) \\ &= \left(\sigma_0^2(Z_0^2, 1), \sigma_0^2 A_1(Z_1^2, 1), \dots, \sigma_{h-1}^2 A_h(Z_h^2, 1) \right) + \mathbf{R}_h = \mathbf{C}_h + \mathbf{R}_h.\end{aligned}$$

Under the assumptions of Theorem 2.1 on Z , each of the random variables σ_t^2 is regularly varying with index $\kappa/2$ and therefore the tail of \mathbf{R}_h is small compared to the tail of $\mathbf{Y}_h^{(2)}$. Hence the tail of $\mathbf{Y}_h^{(2)}$ is determined only by the tail of \mathbf{C}_h . By the same argument and induction we may conclude that the tail of $\mathbf{Y}_h^{(2)}$ is determined by the tail of the vector $\sigma_0^2 \mathbf{Z}_h^{(2)}$. Hence for any Borel set $B \subset \mathbb{S}^{2h-1}$ and in view of Breiman's result (2.9), as $t \rightarrow \infty$,

$$\begin{aligned}\frac{P\left(|\mathbf{Y}_h^{(2)}| > xt, \mathbf{Y}_h^{(2)}/|\mathbf{Y}_h^{(2)}| \in B\right)}{P\left(|\mathbf{Y}_h^{(2)}| > t\right)} &\sim \frac{P\left(\sigma_0^2|\mathbf{Z}_h^{(2)}| > xt, \mathbf{Z}_h^{(2)}/|\mathbf{Z}_h^{(2)}| \in B\right)}{P\left(\sigma_0^2|\mathbf{Z}_h^{(2)}| > t\right)} \\ &\sim x^{-\kappa/2} \frac{E|\mathbf{Z}_h^{(2)}|^{\kappa/2} I_{\{\mathbf{Z}_h^{(2)}/|\mathbf{Z}_h^{(2)}| \in B\}}}{E|\mathbf{Z}_h^{(2)}|^{\kappa/2}} \\ &=: x^{-\kappa/2} P(\Theta \in B).\end{aligned}$$

Since $\mathbf{Y}_h^{(2)}$ is positive with probability 1, it follows from the results in the Appendix of [12] that $\mathbf{Y}_h^{(1)}$ is regularly varying with index κ and spectral measure $P_{\Theta^{1/2}}$. It remains to consider \mathbf{Y}_h . We can write $\mathbf{Y}_h = (\text{sign}(Z_0)|X_0|, \sigma_0, \dots, \text{sign}(Z_h)|X_h|, \sigma_h)$, and, by symmetry of Z , we know that the sequence $(\text{sign}(Z_t))$ is independent of the sequence $(|X_t|, \sigma_t)$. Therefore we can use the results in the Appendix of [12] to conclude that \mathbf{Y}_h is regularly varying with index κ and spectral measure given by the distribution of the vector (2.11). \square

3. Convergence of point processes. We follow the point process theory in Kallenberg [28]. The state space of the point processes considered is $\overline{\mathbb{R}}^d \setminus \{\mathbf{0}\}$. Write \mathcal{M} for the collection of Radon counting measures on $\overline{\mathbb{R}} \setminus \{\mathbf{0}\}$. Define

$$(3.1) \quad \widetilde{\mathcal{M}} = \{\mu \in \mathcal{M} : \mu(\{\mathbf{x} : |\mathbf{x}| > 1\}) = 0 \quad \text{and} \quad \mu(\{\mathbf{x} : \mathbf{x} \in \mathbb{S}^{d-1}\}) > 0\},$$

and let $\mathcal{B}(\widetilde{\mathcal{M}})$ be the Borel σ -field of $\widetilde{\mathcal{M}}$.

Let (X_t) be a strictly stationary GARCH(1, 1) process. For fixed $h \geq 0$, we consider the strictly stationary sequence of random row vectors

$$\mathbf{X}_t = (X_t, \sigma_t, \dots, X_{t+h}, \sigma_{t+h}), \quad t \geq 1.$$

Under the conditions of Theorem 2.3, \mathbf{X} is regularly varying in $\mathbb{R}^{2(h+1)}$ with index $\kappa > 0$, and so are $|\mathbf{X}|$ and X_t in \mathbb{R} . Standard theory for regularly varying functions implies that there exists a sequence (a_n) such that

$$nP(|\mathbf{X}| > a_n) \rightarrow 1, \quad n \rightarrow \infty,$$

and $a_n = n^{1/\kappa} \ell(n)$ for a slowly varying function ℓ . In what follows, $\varepsilon_{\mathbf{x}}$ denotes Dirac measure at \mathbf{x} .

The following theorem is our main result on weak convergence.

THEOREM 3.1. *Let (X_t) be a GARCH(1, 1) process satisfying the conditions of Theorem 2.1(b). Then*

$$N_n := \sum_{t=1}^n \varepsilon_{\mathbf{X}_t/a_n} \xrightarrow{d} N := \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i \mathbf{Q}_{ij}},$$

where \xrightarrow{d} denotes convergence in distribution in $(\mathcal{M}, \mathcal{B}(\mathcal{M}))$, $\sum_{i=1}^{\infty} \varepsilon_{P_i}$ is a Poisson process on \mathbb{R}_+ with intensity measure $\nu(dy) = \theta_X \kappa y^{-\kappa-1} dy$, κ is the solution to (2.7) and θ_X is the extremal index of the sequence $(|\mathbf{X}_t|)$ which exists and is positive; see Section 4 for a definition. The process (P_i) is independent of the sequence of iid point processes $\sum_{j=1}^{\infty} \varepsilon_{\mathbf{Q}_{ij}}$, $i \geq 1$, with common distribution Q on $(\mathcal{M}, \tilde{B}(\mathcal{M}))$. The latter distribution is specified in Theorem 2.8 of [11].

We write

$$\mathbf{Q}_{ij} = ((Q_{ij,X}^{(m)}, Q_{ij,\sigma}^{(m)}), m = 0, \dots, h).$$

PROOF. The theorem is a consequence of Theorem 2.8 in [11]. The latter requires the joint regular variation of all finite-dimensional distributions of (\mathbf{X}_t) , which follows from Theorem 2.3, and a weak mixing condition on (\mathbf{X}_t) , which is trivially satisfied since the process is strongly mixing with geometric rate ϕ_n ; see [12]. Moreover, one needs to check the condition

$$(3.2) \quad \lim_{\kappa \rightarrow \infty} \limsup_{n \rightarrow \infty} P\left(\bigvee_{\kappa \leq |t| \leq r_n} |\mathbf{X}_t| > a_n y \mid |\mathbf{X}_0| > a_n y\right) = 0, \quad y > 0,$$

where $\bigvee_i b_i = \max_i b_i$ and $r_n, m_n \rightarrow \infty$ are two integer sequences such that $n\phi_{m_n}/r_n \rightarrow 0$, $r_n m_n/n \rightarrow 0$. By the definition of the sequence (\mathbf{X}_t) , it suffices to switch in condition (3.2) to the sequence (X_t^2, σ_t^2) and to replace $a_n y$ by $a_n^2 y^2$. Recall that the former sequence satisfies the recurrence equation (2.1). In this situation one can apply the techniques of the proof of Theorem 2.3 in [12] to conclude that (3.2) holds. \square

REMARK 3.2. Analogous results can be obtained for the vectors

$$\mathbf{X}_t^{(l)} = (|X_t|^l, \sigma_t^l, \dots, |X_{t+h}|^l, \sigma_{t+h}^l), \quad l = 1, 2,$$

either by applying the same arguments of proof as above or by deriving the weak limit of the point processes from Theorem 3.1 in combination with a continuous mapping argument. Indeed, under the assumptions of Theorem 3.1,

$$N_n^{(l)} = \sum_{t=1}^n \varepsilon_{\mathbf{X}_t^{(l)}/a_n^l} \xrightarrow{d} N^{(l)} = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i^l \mathbf{Q}_{ij}^{(l)}}, \quad l = 1, 2,$$

where (P_i) is the same as above and

$$\mathbf{Q}_{ij}^{(l)} = \left(\left(\left| Q_{ij,X}^{(m)} \right|^l, \left| Q_{ij,\sigma}^{(m)} \right|^l \right), m = 0, \dots, h \right), \quad l = 1, 2.$$

REMARK 3.3. It is possible to extend the above results to point processes with points in time-space. For illustrational purposes we restrict ourselves to the processes $\widehat{N}_n = \sum_{t=1}^n \varepsilon_{(t/n, X_t/a_n)}$. The analogous results for the point processes $\widehat{N}_n^{(l)}$ for $(|X_t|^l)$ are also valid. The weak convergence of (N_n) implies the convergence of (\widehat{N}_n) under the assumption of strong mixing; see [36]. For fixed $x > 0$, the point process of exceedances of the threshold xa_n by the sequence (X_t) is defined as

$$\widetilde{N}_n(\cdot) = \sum_{t=1}^n \varepsilon_{t/n}(\cdot) I_{\{X_t > xa_n\}} = \widehat{N}_n(\cdot \times (x, \infty)).$$

According to a result in [26] (cf. [18]), the weak limit of (\widetilde{N}_n) is compound Poisson with compounding probabilities π_k and probability generating function $\Pi(u) = \sum_{k=1}^{\infty} \pi_k u^k$. Specifically, in the limiting compound Poisson process events occur as an ordinary Poisson process, independent of the multiplicities (cluster sizes) of the events with compounding probabilities π_k ; that is, π_k is the probability that an event has multiplicity k .

4. The extremal behavior. The point process results of Section 3 enable one to study the extremal behavior of the sequences (X_t) and (σ_t) . We assume that the conditions of Theorem 2.1(b) hold. For any sequence of random variables (Y_n) define the partial maxima

$$M_{n,Y} = \max_{i=1,\dots,n} Y_i, \quad n = 1, 2, \dots$$

Assume that (Y_n) is iid with the same marginal distribution as X . Hence we may conclude that $P(Y > x) \sim cx^{-\kappa}$ for some $\kappa > 0$. Then, with (b_n) such that $nP(X > b_n) \sim 1$,

$$(4.1) \quad P(b_n^{-1} M_{n,Y} \leq x) \rightarrow P(Y^{(\kappa)} \leq x), \quad x > 0,$$

where $Y^{(\kappa)}$ has a standard Fréchet distribution function $\Phi_{\kappa}(x) = \exp(-x^{-\kappa})$, $x > 0$; see for example [14], Chapter 3. For dependent sequences such as (X_t) and (σ_t) exceedances of high thresholds occur in clusters, and so we cannot expect that (4.1) remains valid for them. For a stationary sequence (Y_t) the notion of *extremal index* θ_Y describes the clustering behavior of the extremes (see Leadbetter [30]; cf. [14], Section 8.1):

$$P(b_n^{-1} M_{n,Y} \leq x) \rightarrow [P(Y^{(\kappa)} \leq x)]^{\theta_Y}, \quad x > 0.$$

The extremal index θ_Y assumes values in $[0,1]$ and can be interpreted as the reciprocal of the expected cluster size of high-level exceedances of the normalized sequence (Y_t) . We define the sequence $(b_{n,Y})$ for a stationary sequence

(Y_t) as $nP(Y > b_{n,Y}) \sim 1$. In particular, $b_{n,\sigma}$, $b_{n,|X|}$ and $b_{n,X}$, up to a multiplicative constant, are asymptotically of the same order as $n^{1/\kappa}$. Recall that we write $A_t = \alpha_1 Z_{t-1}^2 + \beta_1$, $t = 1, 2, \dots$.

THEOREM 4.1. *Assume the conditions of Theorem 2.1(b) are satisfied.*

(a) *The partial maxima of (σ_t) satisfy the limit relation*

$$P(b_{n,\sigma}^{-1} M_{n,\sigma} \leq x) \rightarrow [P(Y^{(\kappa)} \leq x)]^{\theta_\sigma}, \quad x > 0,$$

with extremal index

$$\theta_\sigma = \int_1^\infty P\left(\sup_{t \geq 1} \prod_{j=1}^t A_j \leq y^{-1}\right)^{\frac{\kappa}{2}} y^{-(\kappa/2)-1} dy.$$

(b) *The partial maxima of $(|X_t|)$ satisfy the limit relation*

$$(4.2) \quad P(b_{n,|X|}^{-1} M_{n,|X|} \leq x) \rightarrow [P(Y^{(\kappa)} \leq x)]^{\theta_{|X|}}, \quad x > 0,$$

with extremal index

$$\theta_{|X|} = \lim_{k \rightarrow \infty} E\left(|Z_1|^\kappa - \max_{j=2,\dots,k+1} \left|Z_j^2 \prod_{i=2}^j A_i\right|^{\kappa/2}\right)_+ / E|Z_1|^\kappa.$$

(c) *The partial maxima of (X_t) satisfy the limit relation*

$$P(b_{n,X}^{-1} M_{n,X} \leq x) \rightarrow [P(Y^{(\kappa)} \leq x)]^{\theta_X}, \quad x > 0$$

with extremal index $\theta_X = 2\theta_{|X|}(1 - \tilde{\Pi}(0.5))$. Here $\tilde{\Pi}$ is the probability generating function corresponding to the limiting compound Poisson process of the point processes of exceedances of the thresholds $b_{n,|X|}$ by $(|X_t|)$; see Remark 3.3.

The formulas for the extremal indices given above can be evaluated numerically or by Monte Carlo techniques. An example of how to proceed for an ARCH(1) process has been given in [13]. We restrict ourselves to the statistical estimation of these indices for some simulated GARCH(1, 1) and foreign exchange log-returns; see Section 6.

PROOF. (a) Taking into account that $\theta_\sigma = \theta_{\sigma^2}$, the proof follows by an applications of Theorem 2.1 in [13] to the recurrence equation (2.3) for the sequence (σ_t^2) .

(b) The existence of the Frechét limit in (4.2) follows from Theorem 3.3.3 in [31], the fact that (X_t) is strongly mixing and the Pareto-like tails of X . The existence of the extremal index $\theta_{|X|}$ is a consequence of Theorem 2.8 in [11]. For the calculation of $\theta_{|X|} = \theta_{X^2}$ we follow the lines of the calculation of the extremal index of an ARCH(1) in Remark 4.2 of [11]. In the notation of the latter paper, if $\theta = (\theta_{-k}^{(k)}, \dots, \theta_k^{(k)})$ is the $(2k+1)$ -dimensional random row vector

that appears in the definition of the joint regular variation of (X_{-k}^2, \dots, X_k^2) by virtue of Theorem 2.3(a) above,

$$P(\theta \in A) = E \left[|Z_{2k+1}^{(2)}|^{k/2} I \left(\mathbf{Z}_{2k+1}^{(2)} / |\mathbf{Z}_{2k+1}^{(2)}| \in A \right) \right] / E |\mathbf{Z}_{2k+1}^{(2)}|^{k/2},$$

where $\mathbf{Z}_{2k+1}^{(2)} = (Z_0^2, A_1 Z_1^2, \dots, \prod_{j=1}^{2k} A_j Z_{2k}^2)$. Hence, for any measurable function g ,

$$E g(\theta) = E \left[g(\mathbf{Z}_{2k+1}^{(2)} / |\mathbf{Z}_{2k+1}^{(2)}|) |\mathbf{Z}_{2k+1}^{(2)}|^{k/2} \right] / E |\mathbf{Z}_{2k+1}^{(2)}|^{k/2}.$$

Therefore, together with (2.11) of [11], this yields θ_{X^2} as the limit for $k \rightarrow \infty$ of the quantities

$$\begin{aligned} & E \left(\left[\theta_0^{(k)} \right]^{k/2} - \bigvee_{j=1}^k \left[\theta_j^{(k)} \right]^{k/2} \right)_+ / E \left[\theta_0^{(k)} \right]^{k/2} \\ &= E \left(\left| Z_k^2 \prod_{i=1}^k A_i \right|^{k/2} - \max_{j=k+1, \dots, 2k} \left| Z_j^2 \prod_{i=1}^j A_i \right|^{k/2} \right)_+ / E \left| Z_k^2 \prod_{i=1}^k A_i \right|^{k/2} \\ &= E \left(|Z_k|^\kappa - \max_{j=k+1, \dots, 2k} \left| Z_j^2 \prod_{i=k+1}^j A_i \right|^{k/2} \right)_+ / E |Z|^\kappa. \end{aligned}$$

(c) Notice that, by symmetry of Z , $(X_t) = (r_t |X_t|)$, where the sequence of the $r_t = \text{sign}(X_t)$ is independent of $(|X_t|)$. As in [13] for the ARCH(1), one can use this property to obtain the limit distribution of $(b_{n,X}^{-1} M_{n,X})$ and the extremal index θ_X by independent thinning from the point processes of exceedances for $(|X_t|)$. The weak convergence of the latter processes to a compound Poisson process has been described in Remark 3.3. Then proceed as in [13], pages 222–223. \square

5. Convergence of the sample autocorrelations. In this section we study the weak limit behavior of the sample autocovariances and sample autocorrelations of the sequences (X_t) and (σ_t) , their squares and absolute values. We assume that the conditions of Theorem 2.1(b) hold. Then the vector (X_t, σ_t) is regularly varying with index $\kappa > 0$ and, by Theorem 3.1 and Remark 3.2, the point processes $N_n, N_n^{(1)}$ and $N_n^{(2)}$ generated by the vectors $\mathbf{X}_t, \mathbf{X}_t^{(1)}$ and $\mathbf{X}_t^{(2)}$, respectively, converge in distribution to the process $N, N^{(1)}$ and $N^{(2)}$. This is the basis for the weak convergence of the sample autocovariance function (ACVF) $\gamma_{n,X}$ and the sample autocorrelation function (ACF) $\rho_{n,X}$ defined as

$$\gamma_{n,X}(h) = \frac{1}{n} \sum_{t=1}^{n-h} X_t X_{t+h} \quad \text{and} \quad \rho_{n,X} = \gamma_{n,X}(h) / \gamma_{n,X}(0), \quad h = 0, 1, \dots$$

The sample ACVF/ACF for the sequences $(|X_t|), (X_t^2), (\sigma_t)$ and (σ_t^2) are defined analogously. The deterministic counterparts (ACVF, ACF) are

denoted by

$$\begin{aligned}\gamma_X(h) &= EX_0X_h, \quad \gamma_{|X|}(h) = E|X_0X_h|, \quad \gamma_{X^2}(h) = EX_0^2X_h^2, \\ \rho_X(h) &= \gamma_X(h)/\gamma_X(0) \quad \text{etc.}\end{aligned}$$

Usually, centering around the mean-sample mean is included in the definition of the ACF-sample ACF. However, our choice to focus on the uncentered versions is justified by the fact that in the heavy-tailed case centering with the sample mean is not relevant for asymptotic results; similar arguments to the ones below involving the centered versions yield the same limits.

In what follows, we frequently use the notion of multivariate α -stable distribution. We refer to [42] for an encyclopedic treatment of multivariate stable distributions.

5.1. Convergence in distribution of the sample ACF. We first consider the cases when the sample ACFs have nondegenerate distributional limits. Then \mathbf{X}_t is regularly varying with index $\kappa > 0$ and

$$N_n \xrightarrow{d} N = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i \mathbf{Q}_{ij}} \quad \text{and} \quad N_n^{(l)} \xrightarrow{d} N^{(l)}, \quad l = 1, 2.$$

THE CASE $\kappa \in (0, 2)$. An application of Theorem 3.5 in [11] yields

$$\begin{aligned}(na_n^{-2}(\gamma_{n,X}(m), \gamma_{n,\sigma}(m)))_{m=0,\dots,h} &\xrightarrow{d} ((V_{m,X}, V_{m,\sigma}))_{m=0,\dots,h}, \\ ((\rho_{n,X}(m), \rho_{n,\sigma}(m)))_{m=1,\dots,h} &\xrightarrow{d} \left(\left(\frac{V_{m,X}}{V_{0,X}}, \frac{V_{m,\sigma}}{V_{0,\sigma}} \right) \right)_{m=1,\dots,h},\end{aligned}$$

where the vector $((V_{m,X}, V_{m,\sigma}))_{m=1,\dots,h}$ is $\kappa/2$ -stable with point process representation

$$(5.1) \quad V_{m,X} = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i^2 Q_{ij,X}^{(0)} Q_{ij,X}^{(m)}, \quad m = 0, \dots, h,$$

and the $V_{m,\sigma}$'s are defined by replacing in (5.1) $Q_{ij,X}^{(0)} Q_{ij,X}^{(m)}$ with $Q_{ij,\sigma}^{(0)} Q_{ij,\sigma}^{(m)}$. The analogous relations hold for the sample ACVF and ACF of the sequences $(|X_t|)$ and (σ_t) . In this case, one has to replace $((V_{m,X}, V_{m,\sigma}))$ with $((V_{m,|X|}, V_{m,\sigma}))$, where $V_{m,|X|}$ is obtained by replacing the vectors \mathbf{Q}_{ij} in the infinite series (5.1) with $\mathbf{Q}_{ij}^{(1)} = (|Q_{ij,X}^{(m)}|, Q_{ij,\sigma}^{(m)})$, $m = 0, \dots, h$.

THE CASE $\kappa \in (0, 4)$. The same argument as above gives

$$\begin{aligned}(na_n^{-4}(\gamma_{n,X^2}(m), \gamma_{n,\sigma^2}(m)))_{m=0,\dots,h} &\xrightarrow{d} ((V_{m,X^2}, V_{m,\sigma^2}))_{m=0,\dots,h}, \\ ((\rho_{n,X^2}(m), \rho_{n,\sigma^2}(m)))_{m=1,\dots,h} &\xrightarrow{d} \left(\left(\frac{V_{m,X^2}}{V_{0,X^2}}, \frac{V_{m,\sigma^2}}{V_{0,\sigma^2}} \right) \right)_{m=1,\dots,h},\end{aligned}$$

where the vector $((V_{m,X^2}, V_{m,\sigma^2}))_{m=1,\dots,h}$ is $\kappa/4$ -stable.

5.2. *Rates of convergence for the sample ACF toward the ACF.* In this section we assume that the covariances of X_t (respectively X_t^2) are finite. Then, by the ergodic theorem, $\gamma_{n,X}(h) \rightarrow \gamma_X(h)$ a.s., $\gamma_{n|X|^l}(h) \rightarrow \gamma_{|X|^l}(h)$ a.s., $l = 1, 2$, and the analogous relations hold for the σ - and σ^2 -sequences. It brings up the question as to the rate of convergence in these results.

5.2.1. *Convergence to the normal distribution.* The Markov chain (X_t^2, σ_t^2) is strongly mixing with geometric rate; see [12]. Hence the standard CLT for strongly mixing sequences applies provided suitable moment conditions hold; see for example [27, 33].

THE CASE $\kappa \in (8, \infty)$. Then $E[X^8 + \sigma^8] < \infty$. The standard CLT applies to the sample ACVF of the X^2 - and σ^2 -sequences,

$$\begin{aligned} n^{1/2}((\gamma_{n,X^2}(m) - \gamma_{X^2}(m), \gamma_{n,\sigma^2}(m) - \gamma_{\sigma^2}(m)))_{m=0,\dots,h} \\ \xrightarrow{d} ((G_{m,X^2}, G_{m,\sigma^2}))_{m=0,\dots,h}, \end{aligned}$$

where the limit is multivariate Gaussian with mean zero. The CLT for the sample ACF follows by an application of the continuous mapping theorem.

THE CASE $\kappa \in (4, \infty)$. Analogous results hold for the X -, $|X|$ - and σ -sequences. We omit the details.

5.2.2. *Convergence to infinite variance stable distributions.* The derivation of these results can be quite technical. We restrict ourselves to explaining the basic ideas and refer to [11, 34] for more details. The weak limits of the sample ACVF is characterized in terms of limiting point processes; the limit of the sample ACF follows from a simple continuous mapping argument. Weak limits of the vectors of sample autocovariances are infinite variance stable random vectors. They are functionals of point processes. It is, however, difficult to describe the spectral measure of these stable vectors analytically; the latter determines the dependence structure of the vector. See [42] for details. *Therefore the results below are qualitative ones.* It is not clear how to use these results for the construction of asymptotic confidence bands for the sample ACFs and ACVFs. This issue needs further investigation.

THE CASE $\kappa \in (4, 8)$. We commence with the sequences (X_t^2) and (σ_t^2) . We establish joint convergence of the sample ACVF to $\kappa/4$ -stable limits directly from the point process convergence

$$(5.2) \quad N_n^{(2)} \sum_{t=1}^{\infty} \varepsilon_{a_n^{-2}(X_t^2, \sigma_t^2, \dots, X_{t+h}^2, \sigma_{t+h}^2)} \xrightarrow{d} N^{(2)} = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varepsilon_{P_i^2 \mathbf{Q}_{ij}^{(2)}},$$

where $\mathbf{Q}_{ij}^{(2)} = ((|Q_{ij,X}^{(m)}|^2, |Q_{ij,\sigma}^{(m)}|^2), m = 0, \dots, h)$. The weak convergence of the sample ACF is then a straightforward consequence of the continuous mapping theorem.

We start with the σ^2 -sequence and only establish joint convergence of $(\gamma_{n,\sigma^2}(0), \gamma_{n,\sigma^2}(1))$; the extension to arbitrary lags is analogous. Recall that

$A_{t+1} = \alpha_1 Z_t^2 + \beta_1$. Now, using the representation (2.3) and the CLT for $(\sigma_t^2 A_{t+1})$ we obtain

$$\begin{aligned} a_n^{-4} \sum_{t=1}^n (\sigma_{t+1}^4 - E\sigma^4) &= a_n^{-4} \sum_{t=1}^n \left((\alpha_0 + \sigma_t^2 A_{t+1})^2 - E\sigma^4 \right) \\ &= a_n^{-4} \sum_{t=1}^n \sigma_t^4 (A_{t+1}^2 - EA^2) + a_n^{-4} EA^2 \sum_{t=1}^n (\sigma_t^4 - E\sigma^4) + o_P(1). \end{aligned}$$

We conclude that for every $\epsilon > 0$,

$$\begin{aligned} (5.3) \quad & (1 - EA^2)na_n^{-4}(\gamma_{n,\sigma^2}(0) - E\sigma^4) \\ &= a_n^{-4} \sum_{t=1}^n \sigma_t^4 (A_{t+1}^2 - EA^2) I_{\{\sigma_t > a_n \epsilon\}} \\ &+ a_n^{-4} \sum_{t=1}^n \sigma_t^4 (A_{t+1}^2 - EA^2) I_{\{\sigma_t \leq a_n \epsilon\}} + o_P(1) = \text{I} + \text{II} + o_P(1). \end{aligned}$$

Now, by Karamata's theorem (see e.g., [3]) as $n \rightarrow \infty$,

$$(5.4) \quad \text{var}(\text{II}) = \text{const } n(\epsilon a_n)^{-8} E\sigma^8 I_{\{\sigma_t \leq a_n \epsilon\}} \sim \text{const } \epsilon^{8-k} \rightarrow 0 \quad \text{as } \epsilon \rightarrow 0.$$

As for I, let $\mathbf{x}_t = (x_{t,X^2}^{(0)}, x_{t,\sigma^2}^{(0)}, \dots, x_{t,X^2}^{(h)}, x_{t,\sigma^2}^{(h)}) \in \overline{\mathbb{R}}^{2(h+1)} \setminus \{\mathbf{0}\}$ and define the mappings $T_{m,\epsilon,X^2}: \mathcal{M} \rightarrow \overline{\mathbb{R}}$ by

$$\begin{aligned} T_{j,\epsilon,\sigma^2} \left(\sum_{i=1}^{\infty} n_i \epsilon_{\mathbf{x}_i} \right) &= \sum_{i=1}^{\infty} n_i (x_{i,\sigma^2}^{(j)})^2 I_{\{|x_{i,\sigma^2}^{(0)}| > \epsilon\}}, \quad j = 1, 2, \\ T_{m,\epsilon,\sigma^2} \left(\sum_{i=1}^{\infty} n_i \epsilon_{\mathbf{x}_i} \right) &= \sum_{i=1}^{\infty} n_i x_{i,\sigma^2}^{(0)} x_{i,\sigma^2}^{(m-1)} I_{\{|x_{i,\sigma^2}^{(0)}| > \epsilon\}}, \quad m \geq 2. \end{aligned}$$

Since the set $\{\mathbf{x} \in \overline{\mathbb{R}}^{2(h+1)} \setminus \{\mathbf{0}\}: |x^{(l)}| > \epsilon\}$ for any $l \geq 0$ is bounded, the CLT and the convergence in (5.2) imply that

$$\begin{aligned} \text{I} &= a_n^{-4} \sum_{t=1}^n \left((\sigma_t^2 A_{t+1} + \alpha_0)^2 - \sigma_t^4 EA^2 \right) I_{\{\sigma_t > a_n \epsilon\}} + o_P(1) \\ &= a_n^{-4} \sum_{t=1}^n (\sigma_{t+1}^4 - \sigma_t^4 EA^2) I_{\{\sigma_t > a_n \epsilon\}} + o_P(1) \\ &= T_{1,\epsilon,\sigma^2} N_n^{(2)} - EA^2 T_{0,\epsilon,\sigma^2} N_n^{(2)} + o_P(1) \\ &\xrightarrow{d} T_{1,\epsilon,\sigma^2} N^{(2)} - EA^2 T_{0,\epsilon,\sigma^2} N^{(2)} =: S(\epsilon, \infty), \end{aligned}$$

where $ES(\epsilon, \infty) = 0$. Using again (5.4) and the argument in [10], pages 897 and 898, $S(\epsilon, \infty) \xrightarrow{d} V_0^*$, say, as $\epsilon \rightarrow 0$. Turning to (5.3), we finally obtain

$$na_n^{-4}(\gamma_{n,\sigma^2}(0) - E\sigma^4) \xrightarrow{d} [1 - EA^2]^{-1} V_0^* =: V_0.$$

For $\gamma_{n, \sigma^2}(1)$, we proceed as above and write

$$\begin{aligned} na_n^{-4}(\gamma_{n, \sigma^2}(1) - \gamma_{\sigma^2}(1)) &= \alpha_n^{-4} \sum_{t=1}^n (\sigma_t^2(\alpha_0 + \sigma_t^2 A_{t+1}) - \gamma_{\sigma^2}(1)) \\ &= \alpha_n^{-4} \sum_{t=1}^n \sigma_t^4 (A_{t+1} - EA) I_{\{\sigma_t > a_n \varepsilon\}} + \alpha_n^{-4} \sum_{t=1}^n \sigma_t^4 (A_{t+1} - EA) I_{\{\sigma_t \leq a_n \varepsilon\}} \\ &\quad + EA na_n^{-4}(\gamma_{n, \sigma^2}(0) - \gamma_{\sigma^2}(0)) + o_P(1) =: \text{III} + \text{IV} + \text{V} + o_P(1). \end{aligned}$$

As for II, $\lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \text{var}(\text{IV}) = 0$. Moreover,

$$\begin{aligned} \text{III} + \text{V} &= \alpha_n^{-4} \sum_{t=1}^n (\sigma_t^2 \sigma_{t+1}^2 - \sigma_t^4 EA) I_{\{\sigma_t > a_n \varepsilon\}} + EA na_n^{-4}(\gamma_{n, \sigma^2}(0) - \gamma_{\sigma^2}(0)) + o_P(1) \\ &= T_{2, \varepsilon, \sigma^2} N_n^{(2)} - EA T_{1, \varepsilon, \sigma^2} N_n^{(2)} + EA na_n^{-4}(\gamma_{n, \sigma^2}(0) - \gamma_{\sigma^2}(0)) + o_P(1) \\ &\xrightarrow{d} T_{2, \varepsilon, \sigma^2} N^{(2)} - EA T_{1, \varepsilon, \sigma^2} N^{(2)} + EA V_0 \quad \text{as } n \rightarrow \infty \\ &\xrightarrow{d} V_1^* + EA V_0 =: V_1 \quad \text{as } \varepsilon \rightarrow 0. \end{aligned}$$

It also follows from [10], pages 897 and 898, that (V_0, V_1) is jointly $\kappa/4$ -stable.

The weak convergence of $na_n^{-4}(\gamma_{n, X^2}(m) - \gamma_{X^2}(m))$, $m = 0, \dots, h$ follows the same patterns and can indeed be reduced to the convergence of linear combinations of the sample ACVF of the σ^2 -sequence. Notice that the condition $\alpha_1 > 0$ is necessary for the existence of a regularly varying tail for X with index κ provided $E|Z|^\kappa < \infty$; see Remark 2.2. We can write $Z_t^2 = \alpha_1^{-1}((\alpha_1 Z_t^2 + \beta_1) - \beta_1) = \alpha_1^{-1}(A_{t+1} - \beta_1)$, and so, using the CLT, for $m \geq 1$,

$$\begin{aligned} na_n^{-4}(\gamma_{n, X^2}(m) - \gamma_{X^2}(m)) &= \alpha_n^{-4} \alpha_1^{-2} \sum_{t=1}^n \left[\sigma_t^2 \sigma_{t+m}^2 (A_{t+1} - \beta_1)(A_{t+m+1} - \beta_1) - \alpha_1^2 \gamma_{X^2}(m) \right] \\ &= \alpha_n^{-4} \alpha_1^{-2} \sum_{t=1}^n \left[(\sigma_{t+1}^2 \sigma_{t+m+1}^2 - \beta_1 \sigma_t^2 \sigma_{t+m+1}^2 - \beta_1 \sigma_{t+1}^2 \sigma_{t+m}^2 + \beta_1^2 \sigma_t^2 \sigma_{t+m}^2) \right. \\ &\quad \left. - ((1 + \beta_1^2) E \sigma_0^2 \sigma_m^2 - \beta_1 E \sigma_0^2 \sigma_{m+1}^2 - \beta_1 E \sigma_0^2 \sigma_{m-1}^2) \right] + o_P(1) \\ &= na_n^{-4} \alpha_1^{-2} \left[(1 + \beta_1^2)(\gamma_{n, \sigma^2}(m) - \gamma_{\sigma^2}(m)) \right. \\ &\quad \left. - \beta_1(\gamma_{n, \sigma^2}(m+1) - \gamma_{\sigma^2}(m+1)) - \beta_1(\gamma_{n, \sigma^2}(m-1) \right. \\ &\quad \left. - \gamma_{\sigma^2}(m-1)) \right] + o_P(1) \\ &\xrightarrow{d} \alpha_1^{-2} ((1 + \beta_1^2) V_m - \beta_1 V_{m+1} - \beta_1 V_{m-1}). \end{aligned}$$

For $m = 0$ one can get a similar expression for the limit variable. We omit details.

THE CASE $\kappa \in (2, 4)$. Arguments similar to those above show that the joint convergence of a finite number of sample autocovariances from the X -, $|X|$ - and σ -sequences have a multivariate $\kappa/2$ -stable limit. We omit details. The interested reader is referred to [34]. \square

6. An empirical study of simulated data and foreign exchange rates.

In this section we study the sample ACFs of the foreign exchange (FX) rate Japanese yen–U.S. dollar (1992–1996) log-returns and of a fitted GARCH(1, 1) model. Although we focus on one particular series *our findings are typical for log-returns of FX rates, stock indices and share prices*. We consider 70,000 30 minute data (in Olsen's θ -time; see [9] for details). We fitted a GARCH(1, 1) model to the data, using quasi-maximum likelihood estimation (see [22, 21]),

$$(6.1) \quad \alpha_0 = 10^{-7}, \quad \alpha_1 = 0.11, \quad \beta_1 = 0.88.$$

Notice that $\alpha_1 + \beta_1 = 0.99$, a value which is very close to 1. This is a typical situation for various financial time series; see [17, 2, 23]. Given estimates $\hat{\alpha}_1, \hat{\beta}_1$, one can calculate the *residuals* $\hat{Z}_t = X_t/\hat{\sigma}_t$, where $\hat{\sigma}_t^2 = \hat{\alpha}_0 + \hat{\alpha}_1\hat{\sigma}_{t-1}^2 + \hat{\beta}_1X_{t-1}^2$. For notational ease, we will not distinguish between the estimates $\hat{\alpha}_1, \hat{\beta}_1, \hat{Z}_t, \dots$ and the true values $\alpha_1, \beta_1, Z_t, \dots$. Figure 1 displays the QQ-plot of the residuals against the quantiles of the Student (4) distribution with variance 1 (which has tails $P(Z > x) \sim cx^{-4}$), indicating an overall good fit. In [34] we used a least squares method to get an estimate of 3.56 for the tail index of the residuals. In the literature, Student distributions, as heavy-tailed distributions, were fitted to the residuals; see [2]. Given the GARCH(1, 1) model is correct, the choice of a (unit variance) Student distribution for Z is certainly closer to reality than the normal assumption.

The tails of X and σ are determined by the center *and* the tails of the distribution of Z via (2.7). For this reason, in this paper we do not give a precise parametric description of the distribution of the residuals of the FX log-returns. For our purposes, it is more realistic to work with the empirical analogue to (2.7) given by

$$(6.2) \quad \frac{1}{n} \sum_{i=1}^n |\alpha_1 Z_i^2 + \beta_1|^{\kappa/2} = 1.$$

The theoretical basis for this approach is provided by Theorem 1.1 in [38]. There it is shown that, if $E|A|^{\kappa+\delta} < \infty$ for some $\delta > 0$, the solution $\hat{\kappa}$ to (6.2) is asymptotically normal with mean κ and variance $(EA^\kappa - 1)/(nE(A^{\kappa/2} \ln A))$. If the latter condition is not satisfied one can still show consistency of $\hat{\kappa}$ under the assumption $E|A|^{(\kappa/2)+\delta} < \infty$; see [15], Lemma 3.3. Replacing the asymptotic mean and variance by their sample analogues, we obtain the asymptotic 95% confidence band $[2.25 - 0.014, 2.25 + 0.014]$ for κ . (The latter confidence band has to be treated with caution since the finiteness of $E|Z|^{4.5+\delta}$ is questionable

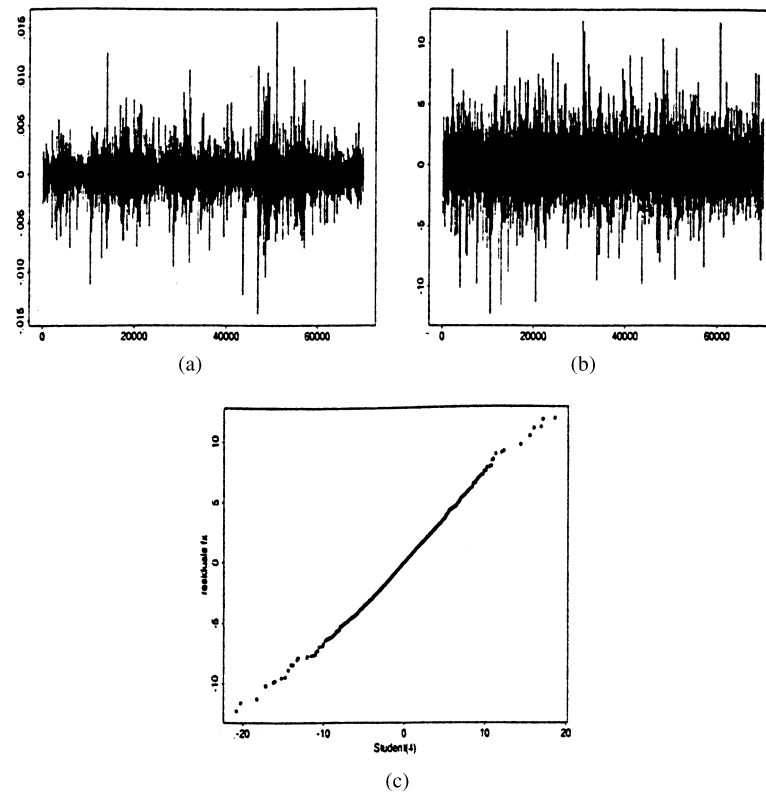


FIG. 1. Plots of the JPY-USD FX log-returns (a) and their residuals (b) from the GARCH(1, 1) model (6.1). (c) QQ-plot of the residuals against the quantiles of a Student(4) distribution with variance 1.

for this data set if one assumes that Z has a Student distribution.) For the GARCH(1, 1) process with parameters (6.1) we thus may conclude under the conditions of Theorem 2.1 that

$$P(\sigma > x) \sim c_0 x^{-2.25} \quad \text{and} \quad P(X > x) \sim E|Z|^{2.25} P(\sigma > x), \quad x \rightarrow \infty.$$

This implies that $EX^2 < \infty$, but $EX^4 = \infty$.

Figure 2 shows the sample ACFs of the FX log-returns and their powers. Confidence bands were derived from 1000 independent repetitions of 70,000 realizations from the GARCH(1, 1) model (6.1). The iid noise was generated from the empirical distribution of the residuals of the FX log-returns. The interpretation of these sample ACFs very much depends on how heavy the tails of the X_t 's are. Using the above findings of a Pareto-like tail for X with $\kappa = 2.25$ and the theory of Section 5, we conclude that the limit distribution of the sample ACFs of the X_t 's and $|X_t|$'s have infinite variance 2.25/2-stable limits with rate of convergence $(na_n^{-2})^{-1} \sim cn^{-1+\kappa/2} = cn^{0.125}$. Notice that $n^{0.125} = 70,000^{0.125} = 4.03$. Thus, despite the large sample size, the asymptotic confidence bands for $\rho_{n,X}(h)$ and $\rho_{n,|X|}(h)$ are huge. This observation is supported by the bands in Figure 2. The slow rate of convergence of these estimators in combination with the extremely heavy tails of the limit distribution raises serious questions about the meaning and quality of these estimators. This remark applies even more to the sample ACFs of the squares and third powers. In those cases, both $\rho_{n,X^2}(h)$ and $\rho_{n,X^3}(h)$ converge in distribution; that is, these statistics do not estimate anything.

The sample ACFs at the first 50 lags, say, of the absolute values of the FX log-returns do not fall within the 95% confidence bands for the corresponding sample ACFs of the GARCH(1, 1) process; see Figure 2. This means that, even when accounting for the statistical uncertainty, the GARCH(1, 1) model does not describe the second-order dependence structure of the FX log-returns sufficiently accurately. On the other hand, the corresponding sample ACFs for the residuals (Figure 3) behave very much like the sample ACF of a finite variance iid sequence or of a moving average process with very small parameters. This has also been observed in [32]. This compliance with the theoretical requirements of the model is a remarkable feature of the GARCH(1, 1) process and contributed greatly to its success. As a conclusion, the GARCH(1, 1) process cannot explain the long-range dependence effect observed in the sample ACFs of the FX log-returns; see Figure 2. Even if we take into account that the sample ACFs of the squares and third powers are not meaningful, the sample ACF of the absolute values, despite its big statistical uncertainty, should decay to zero roughly at an exponential rate, due to the strong mixing property with geometry rate.

Figure 4 displays the results for estimating the extremal indices θ_X of the JPY-USD FX log-returns, θ_Z of their residuals and θ_σ of the volatility sequence. The estimates are based on the so-called *blocks methods* which involves a sufficiently large number of order statistics; see [24, 25, 43]; [13], Section 8.1. Under general conditions, $\hat{\theta}_n$ is consistent and asymptotically normal. The 95% confidence bands and the median were obtained from 400

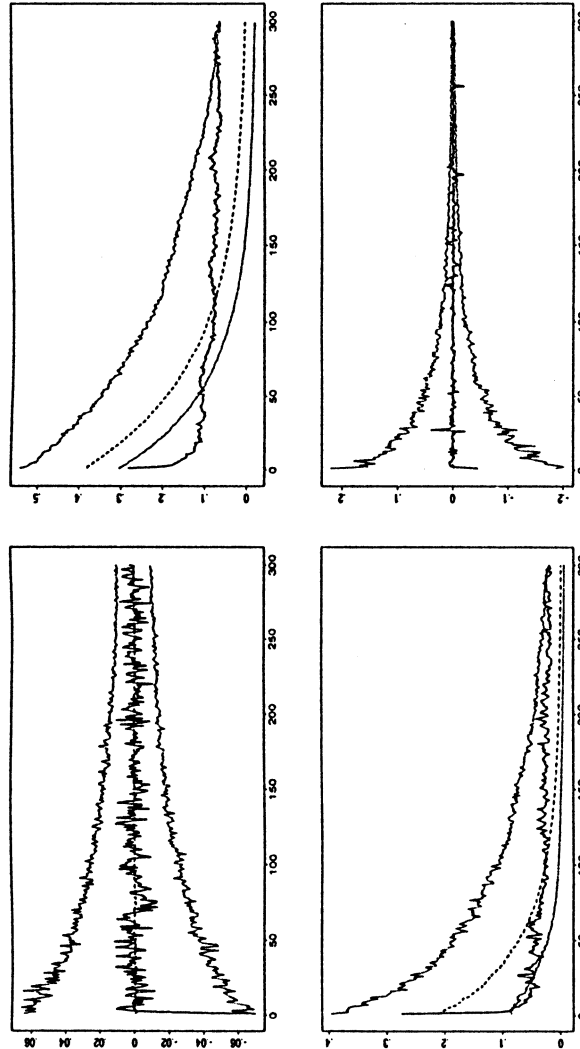


FIG. 2. The sample ACF of the FX log-returns (top, left), their absolute values (top, right), squares (bottom, left) and third powers. The upper and lower solid curves indicate the 2.5% and 97.5%-quantiles of the distributions of the sample ACFs at a fixed lag. The dotted line corresponds to the median of those distributions. In the top left and bottom right figures, the median curve and the sample ACF curve are almost indistinguishable.

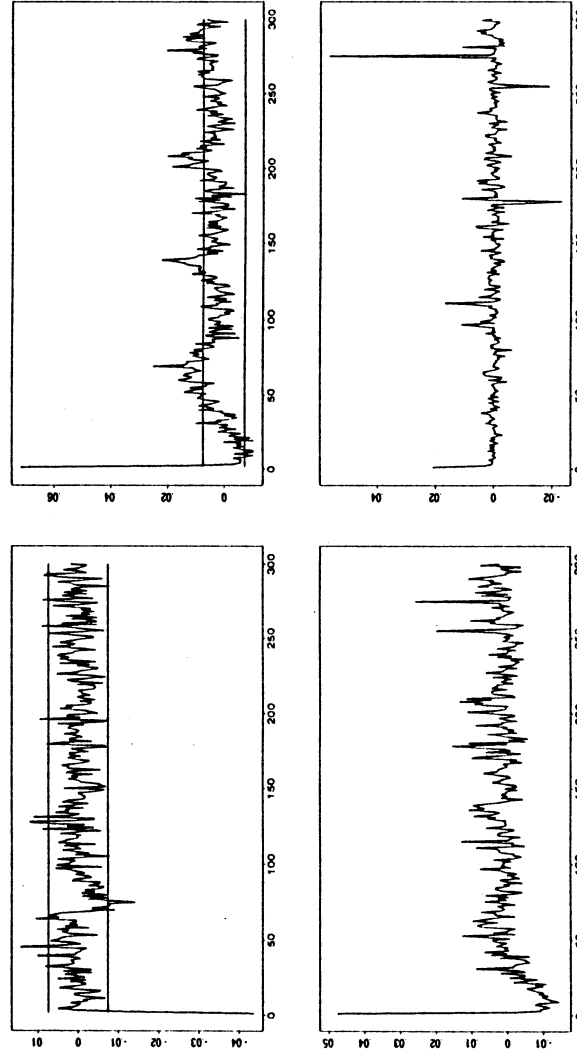


FIG. 3. The sample ACFs of the residuals of the FX log-returns (top, left), their absolute values (top, right), squares (bottom, left) and third powers. The straight lines in the two upper graphs indicate the $\pm 1.96/\sqrt{n}$ asymptotic confidence bands for an iid sequence with finite second moment. In the lower two graphs we refrain from giving \sqrt{n} -confidence bands because Z possibly has an infinite fourth moment. Compare with the sample ACFs of the FX log-returns; see Figure 2. Mind the difference in scale!

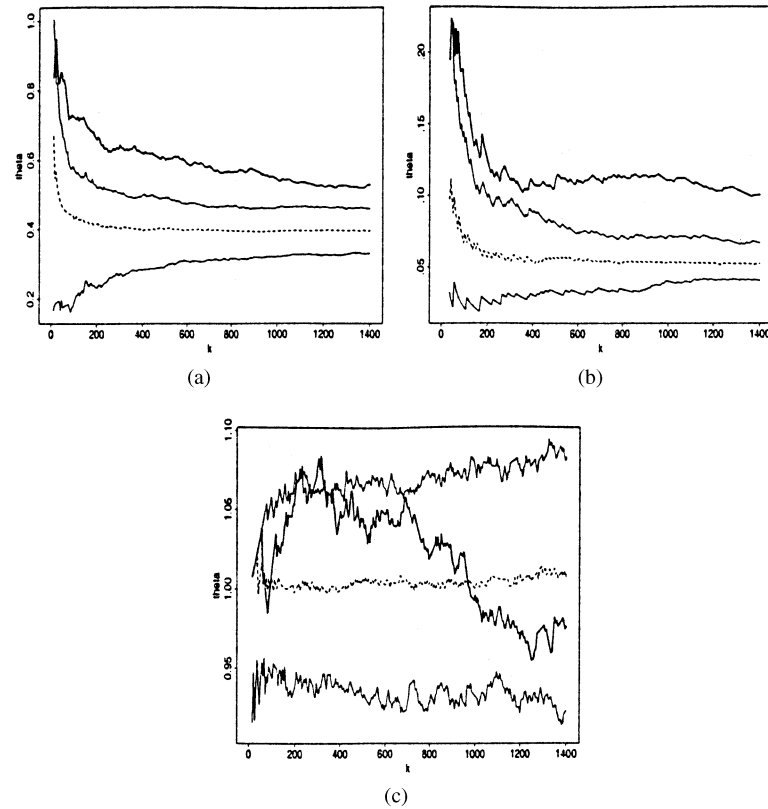


FIG. 4. Estimates of extremal indices as a function of k upper order statistics together with 95% confidence bands and the median (dotted line) based on model (6.1). The estimates of θ_X [(a) upper curve] are above the confidence bands. The same applies for the estimates of θ_σ (b). The estimates of θ_Z (c) are within the confidence bounds.

independent repetitions based on 70,000 GARCH(1, 1) realizations with parameters (6.1) and, as before, the noise was drawn from the empirical distribution of the FX log-return residuals. See [34] for an exact description of the procedure.

The estimates for θ_X and θ_σ of the FX log-returns lie above the 97.5% curve. Hence the expected cluster size is smaller than for the corresponding GARCH(1, 1) model, indicating that *there is less dependence in the tails of the returns series than in the GARCH(1, 1) model*. The corresponding estimates for θ_Z lie within the 95% bands for an iid sequence. This again seems to imply that the residuals very much behave like an iid sequence (with extremal index 1).

Acknowledgments. Both authors thank the Departments of Mathematics and Statistics at Chalmers University and the University of Groningen for financial support. We thank Holger Rootzén for stimulating discussions and Olsen & Associates (Zürich) for making the FX data available to us. Finally, we thank both referees, whose insightful comments led to a substantial improvement of the paper.

REFERENCES

- [1] BABILLOT, M., BOUGEROL, P. and ELIE, L. (1997). The random difference equation $X_n = A_n X_{n-1} + B_n$ in the critical case. *Ann. Probab.* **25** 478–493.
- [2] BAILLIE, R. T. and BOLLERSLEV, T. (1989). The message in daily exchange rates: a conditional-variance tale. *J. Bus. Econom. Statist.* **7** 297–304.
- [3] BINGHAM, N. H., GOLDIE, C. M. and TEUGELS, J. L. (1987). *Regular Variation*. Cambridge Univ. Press.
- [4] BOLLERSLEV, T., CHOU, R. Y. and KRONER, K. F. (1992). ARCH models in finance: a review of the theory and evidence. *J. Econometrics* **52** 5–59.
- [5] BOUGEROL, P. and PICARD, N. (1992). Strict stationarity of generalized autoregressive processes. *Ann. Probab.* **20** 1714–1730.
- [6] BOUGEROL, P. and PICARD, N. (1992). Stationarity of GARCH processes and of some non-negative time series. *J. Econometrics* **52** 115–127.
- [7] BRANDT, A. (1986). The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Adv. Appl. Probab.* **18** 211–220.
- [8] BREIMAN, L. (1965). On some limit theorems similar to the arc sin law. *Theory Probab. Appl.* **10** 323–331.
- [9] DACOROGNA, U. A., MÜLLER, U. A., NAGLER, R. J., OLSEN, R. B. and PICTET, O. V. (1993). A geographical model for the daily and weekly seasonal volatility in the foreign exchange market. *J. Internat. Money and Finance* **12** 413–438.
- [10] DAVIS, R. A. and HSING, T. (1995). Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Ann. Probab.* **23** 879–917.
- [11] DAVIS, R. A. and MIKOSCH, T. (1998). The sample autocorrelations of heavy-tailed processes with applications to ARCH. *Ann. Statist.* **26** 2049–2080.
- [12] DAVIS, R. A., MIKOSCH, T. and BASRAK, B. (1999). The sample ACF of solutions to multivariate stochastic recurrence equations. Preprint. Available at www.math.rug.nl/~mikosch.
- [13] DE HAAN, L., RESNICK, S. I., ROOTZÉN, H. and DE VRIES, C. G. (1989). Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. *Stochastic Process Appl.* **32** 213–224.
- [14] EMBRECHTS, P., KLÜPPELBERG, C. and MIKOSCH, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin.

- [15] EMBRECHTS, P. and MIKOSCH, T. (1991). A bootstrap procedure for estimating the adjustment coefficient. *Insurance Math. Econom.* **10** 181–190.
- [16] ENGLE, R. F. (ed.) (1995). *ARCH Selected Readings*. Oxford Univ. Press.
- [17] ENGLE, R. F. and BOLLERSLEV, T. (1986). Modelling the persistence of conditional variances. (With comments and reply by authors). *Econometric Rev.* **5** 1–87.
- [18] FALK, M., HÜSLER, J. and REISS, R.-D. (1994). *Laws of Small Numbers: Extremes and Rare Events*. Birkhäuser, Basel.
- [19] FURSTENBERG, H. and KESTEN, H. (1960). Products of random matrices. *Ann. Math. Statist.* **31** 457–469.
- [20] GOLDIE, C. M. (1991). Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.* **1** 126–166.
- [21] GOURIEROUX, C. (1997). *ARCH Models and Financial Applications*. Springer, New York.
- [22] GOURIEROUX, C., MONFORT, A. and TROGNON, A. (1984). Pseudo maximum likelihood methods: theory. *Econometrica* **52** 681–700.
- [23] GUILLAUME, D. M., PICTET O. V. and DACOROGNA, M. M. (1995). On the intra-daily performance of GARCH processes. Olsen & Associates (Zürich). Internal Document DMG. 1994-07-31.
- [24] HSING, T. (1991). Estimating the parameters of rare events. *Stochastic Process. Appl.* **37** 117–139.
- [25] HSING, T. (1993). On some estimates based on sample behavior near high level excursions. *Probab. Theory Related Fields* **95** 331–356.
- [26] HSING, T., HÜSLER, J. and LEADBETTER, M. R. (1988). On the exceedance point process for a stationary sequence. *Probab. Theory Related Fields* **78** 97–112.
- [27] IBRAGIMOV, I. A. and LINNIK, YU. V. (1971). *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff, Groningen.
- [28] KALLENBERG, O. (1983). *Random Measures*, 3rd ed. Akademie, Berlin.
- [29] KESTEN, H. (1973). Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131** 207–248.
- [30] LEADBETTER, M. R. (1983). Extremes and local dependence of stationary sequences. *Z. Wahrsch. Verw. Gebiete* **65** 291–306.
- [31] LEADBETTER, M. R., LINDGREN, G. and ROOTZÉN H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer, Berlin.
- [32] LI, W. K. and MAK, T. K. (1994). On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity. *J. Time Ser. Anal.* **15** 627–636.
- [33] MEYN, S. P. and TWEEDIE, R. L. (1993). *Markov Chains and Stochastic Stability*. Springer, London.
- [34] MIKOSCH, T. and STĂRICĂ, C. (1998). Limit theory for the sample autocorrelations and extremes of a GARCH(1, 1) process. Preprint. Available at www.math.rug.nl/~mikosch.
- [35] MIKOSCH, T. and STĂRICĂ, C. (1999). Change of structure in financial time series, long range dependence and the GARCH model. Preprint. Available via www.math.rug.nl/~mikosch.
- [36] MORI, T. (1977). Limit distributions of two-dimensional point processes generated by strong-mixing sequences. *Yokohama Math. J.* **25** 155–168.
- [37] NELSON, D. B. (1990). Stationarity and persistence in the GARCH(1, 1) model. *Econometric Theory* **6** 318–334.
- [38] PITTS, S. M., GRÜBEL, R. and EMBRECHTS, P. (1996). Confidence bounds for the adjustment coefficient. *Adv. in Appl. Probab.* **28** 802–827.
- [39] RESNICK, S. I. (1986). Point processes, regular variation and weak convergence. *Adv. Appl. Probab.* **18** 66–138.
- [40] RESNICK, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer, New York.
- [41] RESNICK, S. I. (1997). Heavy tail modeling and teletraffic data. (With discussion and rejoinder by author). *Ann. Statist.* **25** 1805–1869.
- [42] SAMORODNITSKY, G. and TAQQU, M. S. (1994). *Stable Non-Gaussian Random processes. Stochastic Models with Infinite Variance*. Chapman and Hall, London.

- [43] SMITH, R. L. and WEISSMAN, I. (1994). Estimating the extremal index. *J. Roy. Statist. Soc. Ser. B* **56** 515–528.
- [44] STĂRICA, C and PICTET, O. (1997). The tales the tails of ARCH processes tell. Technical report. Chalmers Univ., Gothenberg.
- [45] VERVAAT, W. (1979). On a stochastic difference equation and a representation of nonnegative infinitely divisible random variables. *Adv. Appl. Probab.* **11** 750–783.
- [46] WALD, A. (1947). *Sequential Analysis*. Wiley, New York.

DEPARTMENT OF MATHEMATICS
P.O. BOX 800
UNIVERSITY OF GRONINGEN
NL-9700 AV GRONINGEN
THE NETHERLANDS
E-MAIL: mikosch@math.rug.nl

DEPARTMENT OF MATHEMATICAL STATISTICS
CHALMERS UNIVERSITY OF TECHNOLOGY
S-412 96 GOTHENBERG
SWEDEN